THE
MILBANK QUARTERLY
A MULTIDISCIPLINARY JOURNAL OF POPULATION HEALTH AND HEALTH POLICY

Evidence on the Efficacy of Inpatient Spending on Medicare Patients

ROBERT KAESTNER and JEFFREY H. SILBER

University of Illinois; University of Pennsylvania

Context: It is widely believed that a significant amount, perhaps as much as 20 to 30 percent, of health care spending in the United States is wasted, despite market forces such as managed care organizations and large, self-insured firms with a financial incentive to eliminate waste of this magnitude.

Methods: This article uses Medicare claims data to study the association between inpatient spending and the thirty-day mortality of Medicare patients admitted to hospitals between 2001 and 2005 for surgery (general, orthopedic, vascular) and medical conditions (acute myocardial infarction [AMI], congestive heart failure [CHF], stroke, and gastrointestinal bleeding).

Findings: Estimates from the analysis indicated that except for AMI patients, a 10 percent increase in inpatient spending was associated with a decrease of between 3.1 and 11.3 percent in thirty-day mortality, depending on the type of patient.

Conclusions: Although some spending may be inefficient, the results suggest that the amount of waste is less than conventionally believed, at least for inpatient care.

Keywords: Efficiency, inpatient spending, mortality.

The efficacy of medical care and medical care spending in producing health has been a topic of central concern to health economics since the inception of the field in the 1960s. From one of the first systematic analyses of the issue, Fuchs concluded: "Although

---

Address correspondence to: Robert Kaestner, Institute of Government and Public Affairs, University of Illinois, 815 West Van Buren Street, Suite 525, Chicago, IL 60614 (email: kaestner@uic.edu).

The Milbank Quarterly, Vol. 88, No. 4, 2010 (pp. 560–594)

many health services definitely improve health, in other cases even the best known techniques may have no effect. This problem of relating input to output is one of the most difficult ones facing economists who try to do research on the health industry" (Fuchs 1966, p. 80). Fuchs's characterization of the contribution of health services to health would now be referred to as an example of "flat-of-the-curve" medicine. His description is noteworthy, because at that time the United States was spending less than $230 per person on health care and health care spending represented just 6 percent of the nation's output. Now, however, we spend approximately $8,290 per person, and health care spending represents nearly 17.3 percent of national income.[1] The same "flat-of-the-curve" argument was used to characterize medical spending during the 1970s. Gruber summarized the findings from the Rand Health Insurance Experiment, which took place between roughly 1975 and 1980, as follows: "It suggests that, at least at the time of the experiment, the typical enrollee in the study was on the 'flat of the medical effectiveness curve,' the portion where additional care was not buying medically effective care" (Gruber 2006, 5).

Fuchs returned to the issue some forty years later, stating with respect to spending on medical care for those on Medicare: "The bottom line is that a considerable amount of the care delivered in the United States is 'flat-of-the-curve' medicine" (Fuchs 2004, 105).

These conclusions by well-known health economists over the past forty years reflect the widely held view that for a long time, a significant amount of health care spending in the United States was unnecessary and wasteful. Indeed, this view has seeped into mainstream media, as exemplified by this quotation from *Consumer Reports*: "Though the idea that more health care is better seems to make sense, recent research has shown that none of the above [specialist care] necessarily helps you live better or longer. In fact, too much medical care might shorten your life" (Consumer Reports Health.org 2008).

Moreover, the amount of supposed waste is thought to be quite large. According to some estimates, the United States is not at the beginning, but well along the "flat-of-the-curve." For example, research by those associated with the *Dartmouth Atlas of Health Care* concluded that Medicare spending in recent years could be reduced by approximately 20 to 30 percent with no impact on mortality (Fisher et al. 2003; Skinner, Fisher, and Wennberg 2005). The inefficiency of medical care spending also is not thought to be limited to government programs. That is, the results of the Rand Health Insurance Experiment suggested that in the late 1970s, 30 percent of spending in the private insurance market had no noticeable effect on health (Newhouse 1993). Finally, the United States does in fact spend nearly twice as much on health care per capita than do most other developed countries but has little to show for it in terms of life expectancy.

While the "flat-of-the-curve" argument is conventional wisdom and appears to have some empirical validity, there are reasons to be skeptical of the supposed magnitude and persistence of the problem. First, market forces should eliminate this inefficiency. Although organizations such as large corporations that self-insure medical expenditures, large health insurers in competitive markets, and managed care organizations participating in publicly financed programs such as Medicare and Medicaid have an incentive to eliminate inefficient health care spending, the available evidence suggests that these organizations have not rooted out much of the supposed inefficiency.[2] The failure of these market forces to eliminate such large amounts of inefficiency is particularly surprising, given that many people believe that the cause of the problem is relatively easy to identify. For example, Fisher and colleagues (2009) argued that much of the inefficiency in health care spending stems from differences in the way that physicians treat patients for illnesses for which the proper course of treatment is uncertain ("gray areas"). That is, some physicians treat patients aggressively and use many resources that have little benefit. According to these scholars, solving the problem is straightforward: using capitated and bundled payments that encourage integrated systems of care and restrain the use of unproven treatments. The mystery is why these changes have not been implemented by managed care organizations, large insurers, and self-insured firms, for all these organizations have the expertise and financial incentive to eliminate such low-value care.

Skinner and Staiger noted a similar puzzle for why hospitals have not adopted low-cost, highly effective "technologies" such as providing aspirin, beta blockers, and reperfusion for heart attack patients. "The real puzzle is why many physicians and hospitals remain so far behind the production possibility frontier, contributing to a remarkable degree of productive inefficiency in health care" (Skinner and Staiger 2009, 4). We would add that the puzzle noted by Skinner and Staiger is particularly perplexing because of the failure of insurers, self-insured firms, and managed care organizations to steer patients away from these lagging physicians and hospitals. These organizations have the ability (e.g., access to large amounts of data), expertise, and financial incentive to identify and eliminate the inefficiencies noted by Skinner and Staiger.

A second reason to be skeptical of the "flat-of-the-curve" hypothesis is that much of the evidence supporting it comes from nonexperimental studies, with the notable exception of the Rand Health Insurance Experiment. The most obvious problem with nonexperimental research in this area is that causality is just as likely to go from health to spending as it is from spending to health, since more resources are typically spent on sicker people. Though obvious, this is a difficult problem to overcome empirically. But several recent studies that used credible research designs to address this empirical problem found evidence that additional spending on medical care is actually very effective (Card, Dobkin, and Maestas 2008; Chandra and Staiger 2007; Doyle 2005, 2007).[3]

In summary, the market mechanisms that would tend to eliminate wasteful spending, particularly given the supposed magnitude of the problem, and the few studies using credible research designs suggest that additional study of the "flat-of-the-curve" hypothesis would be useful. In this article, we examine the effect of inpatient spending on the thirty-day mortality of Medicare patients admitted to the hospital between 2001 and 2005.[4] We consider a broad range of patients, including those admitted for surgery (general, orthopedic, vascular) and several medical conditions (acute myocardial infarction [AMI], congestive heart failure [CHF], stroke, and gastrointestinal bleeding), and we analyze the efficacy of spending separately for each outcome.

The difference between our research question and other research on this topic is important. For example, Fisher and colleagues (2003) and Skinner, Fisher, and Wennberg (2005) investigated the association between all Medicare spending and five- and one-year mortality rates. By contrast, we focus only on Medicare spending in the hospital for specific patients. Although ours is a narrower question than, for example, that of Fisher and colleagues (2003), it is clearly related. Inpatient spending accounts for approximately one-third of all health care spending. Therefore, the efficacy of inpatient spending in terms of patients' outcomes is an important part of the larger question of the efficacy of all health care (Medicare) spending.

To address the potential reverse causality—worse health causes greater spending—that likely characterizes the relationship between spending and health outcomes, we used an instrumental variables procedure. That is, we used the evidence from the *Dartmouth Atlas for Health Care* showing that the intensity of treatment and use of resources for patients in a hospital is strongly associated with the intensity of treatment for patients at the end of life in that same hospital. Accordingly, we used these end-of-life measures of treatment of decedents in particular hospitals to predict inpatient spending for patients in those hospitals, and we used the predicted variation among hospitals in inpatient spending to measure the efficacy of medical spending. This procedure is similar to that used by Skinner, Fisher, and Wennberg (2005). The identifying assumption is that the variation among hospitals in end-of-life spending on decedents who had several chronic conditions is not correlated with unmeasured differences among hospitals regarding their patients' health. We provide evidence to support this assumption.

To bolster the validity of the instrumental variables approach, we included an unusually large set of patient-level risk adjusters to minimize the reverse causality problem. We used six months of earlier information about the patients to construct measures of severity, which are strong predictors of thirty-day mortality. In addition, we limited our analysis of surgical patients to those who experienced an in-hospital health shock or complication. Those who suffered from a complication during surgery are a group of particularly (equally) sick patients. These patients have a much higher mortality rate and are more uniformly ill. Moreover, to the extent that most life-threatening health problems stem from the complication, it can be considered a health shock caused by external forces rather than an underlying health risk. Our limiting our sample in this way reflects previous research that used end-of-life measures of spending, which were calculated for a group of particularly (equally) sick patients, as a plausibly more exogenous measure of spending.

The results of our analysis indicate that spending is significantly and negatively associated with thirty-day mortality. The increase in survival was nontrivial: a 10 percent increase in inpatient spending was associated with a 3 to 11 percent increase in survival. We also found significant heterogeneity in the association between spending and mortality by type of hospital. That is, the associations between spending and mortality were much smaller (less negative) in teaching-intensive hospitals.

Research Design

Our objective was to obtain estimates of the association between in-patient spending and the thirty-day mortality of patients covered by

Medicare who were admitted to hospitals for surgery (general, orthopedic and vascular) and medical conditions (AMI, CHF, stroke, and GI bleeding) between 2001 and 2005. We did this using ordinary least squares regression methods.[5] We limited our sample of surgical patients to those who experienced an in-hospital health shock, in order to obtain a more homogenous group of surgical patients in regard to severity of illness. The use of thirty-day mortality (death within thirty days from admission) is the window of observation most commonly used in Medicare studies of hospital quality, because longer follow-ups increase the likelihood that factors not relevant to care delivered at the hospital will confound estimates.[6]

Specifically, we estimated a model such as the following:

$$
\begin{aligned}
&MORT_{ijt} = \beta IN\_SPEND_j + \alpha_t + \sum_k \delta_k Z_{kjt} + X_i\Gamma + v_{ijt} \\
&\mathrm{i} = 1, \ldots, N \qquad \text{(patients)} \\
&\mathrm{j} = 1, \ldots, J \qquad \text{(hospitals)} \\
&\mathrm{t} = 2001, \ldots, 2005 \ \text{(years)}
\end{aligned}
\tag{1}
$$

In equation (1), *MORT* is the indicator of the thirty-day mortality of patient i who was admitted to hospital j in year t for a specific condition, for example, congestive heart failure; *IN_SPEND* is the total inpatient charges tallied by the hospital (not just reimbursed charges) associated with the admission of patient i; Z is hospital characteristics (k = 1 to K) such as bed size, nurse-to-bed ratio, nurse mix, resident-to-bed ratio (a measure of teaching intensity), and average spending on other types of patients in that hospital;[7] and X is an extensive set of patient-level controls for the severity of illness that were constructed from six months of claims data on patients before the index admission. The model also contains year dummy variables (α_t).

As noted earlier, we used an instrumental variables approach to address the likely reverse causality between inpatient spending and mortality. The instrumental variables procedure uses the firmly established relationship between hospitals' end-of-life spending (EOL) or resource use, such as intensive care unit (ICU) days, on decedents and broader measures of hospitals' treatment intensity. EOL spending in hospital referral regions has been shown to be correlated with spending on acute care episodes, use of specialists, ICU days, and total per-capita Medicare payments (*Dartmouth Atlas of Health Care* 2008). In addition, regional

EOL spending has been shown not to be related to regional illness levels (Fisher et al. 2003; Skinner, Fisher, and Wennberg 2005). Therefore, EOL measures of resource use are a plausibly exogenous source of variation in inpatient spending.[8]

The instrumental variables procedure predicts inpatient spending (*IN_SPEND*) in equation (1) using the following regression:

$$IN\_SPEND_{ijt} = \tilde{\beta} EOL_j + \tilde{\alpha}_t + \sum_k \tilde{\delta}_k Z_{kjt} + X_i \tilde{\Gamma} + u_{ijt} \quad (2)$$

Equation (2) contains the same variables as in equation (1), with the addition of end-of-life resource use (*EOL*) instruments, which in our case includes ICU days per decedent, non-ICU hospital days per decedent, and total inpatient spending per decedent. We used various combinations of these instruments to assess the sensitivity of estimates to the choice of instruments. In equation (2), we used the symbol (~) to indicate that the parameters in equations (1) and (2) differed. Note that the EOL measures vary only by hospital, and not by year. Because data on EOL measures are not available by hospital and year, we could not include hospitals' fixed effects in the model. The assumption underlying the instrumental variables approach is that conditional on inpatient spending (and other measured controls), the end-of-life measures of resource use do not affect mortality. Evidence that this approach is reasonable can be found in the appendix.

As noted, our instrumental variables strategy has already been used (Skinner, Fisher, and Wennberg 2005) in studies estimating the reduced form model obtained by substituting equation (2) into equation (1), which yields:

$$MORT_{ijt} = \dddot{\beta} EOL_j + \dddot{\alpha}_t + \sum_k \dddot{\delta}_k Z_{kjt} + X_i \dddot{\Gamma} + e_{ijt} \quad (3)$$

Equation (3), which uses the symbol (⋯) to indicate different parameters, directly relates mortality to end-of-life measures. Earlier studies that estimated a model similar to equation (3) reported that EOL measures were not related to mortality ($\dddot{\beta} = 0$), which implies that the United States practices "flat-of-the-curve" medicine (Fisher et al. 2003; Skinner, Fisher, and Wennberg 2005). With one exception, the thirty-day mortality for AMI patients, we did not find this to be true for the outcomes we studied. Rather, we found a significant and negative effect of EOL resource use on thirty-day mortality ($\ddot{\beta} < 0$), which partly validates the instrumental variables strategy. Tables A3 and A4 in the appendix show these results.

There are a few possible explanations of the differences between our study and earlier studies that estimated the reduced form model given by equation (3). First, we used more recent data, and the results may be sensitive to the time period. Second, while EOL spending is plausibly exogenous, the possibility of reverse causality remains: sicker patients may be found in hospitals that treat more aggressively. Previous studies that used the EOL measures include minimal risk adjustment for patient characteristics and have been implemented primarily at the regional level. In contrast, we used individual-level data and many patient-level controls drawn from the previous six months of claims data. Although these controls modestly affected the estimates of the effect of spending for some outcomes, it was mostly the hospital-specific controls that were the conditioning variables that mattered most in our analyses (the tables present this evidence). Finally, we used different outcomes—thirty-day mortality for several specific illnesses (e.g., CHF) and certain surgical patients—and focused on inpatient spending. Previous studies that used the end-of-life measures examined all spending and one-year or five-year mortality.

A novel aspect of our study was examining the relationship between inpatient spending and thirty-day mortality for surgical patients who experienced an in-hospital complication. Focusing on surgical patients who had a complication yielded a more homogenous sample in regard to severity of illness. First, patients who have surgery and develop complications are far sicker than those who never develop complications (see Silber et al. 1992, 2005, 2007). Second, studies for which physiological data were available concluded that complications are highly dependent on the severity of the patient's illness at admission. In contrast, samples limited to patients who experience an in-hospital complication do not display as large or significant a variation in measures of patient severity (Silber et al. 1992).[9]

Selecting the sample on the basis of having experienced a complication could bias reduced form and instrumental variables estimates if EOL resource use, the instrument, is correlated with the probability of a complication. We investigated this possibility directly (the results are presented in the article in *Health Services Research* by Silber and colleagues 2010) and found that regression estimates of the association between EOL measures and the probability of having a complication were very small, clinically not important, and not always the same sign.[10] For example, an additional ICU day for decedents was associated with (1) a 0.25 percentage point (0.5%) increase in the probability that a patient admitted for general surgery will have a complication; (2) a 0.27 percentage point (0.4%) increase in the probability that a patient admitted for vascular surgery will have a complication; and (3) a 0.20 percentage point (0.6%) decrease in the probability that a patient admitted for orthopedic surgery will experience a complication.

The extremely small magnitudes and mixed signs of these estimates suggest that selecting the surgical sample on the basis of whether a surgical patient experienced a complication does not significantly bias estimates. More important, all surgical patients who die were included in the sample of surgical patients with a complication. In the few instances in which the deaths were not preceded by a complication (identified by claims data), we assumed that the death was caused by an undocumented complication (Silber et al. 2007). Thus, the consequence of using all surgical patients is attenuated estimates of the effect of interest, that is, the effect of spending on saving a life after an in-hospital health shock.

Another type of selection that may be related to the EOL measures is the probability of being admitted to the hospital. Although this may not be particularly problematic for outcomes such as AMI and stroke, for other outcomes including the surgical outcomes, hospital "aggressiveness" as measured by EOL measures could be related to the probability of admission. Here, too, limiting the sample to surgical patients with a complication, which reduces the heterogeneity of patients' severity of illness, arguably eliminates the type of selection associated with the initial admission. Thus, the selection into the hospital was unlikely to be a serious problem in our analysis and was limited to, at most, two outcomes, CHF and GI bleeding. As we show, however, except for AMI, estimates of the association between spending and mortality were quite similar across outcomes, which suggests that the selection of this type was not the cause of the associations we found between inpatient spending and mortality.

Finally, to close the door a little more on this potential selection problem, we divided our sample of surgical patients into three groups based on the variations in surgical admission rates across hospital referral regions (HRRs). For each ICD-9 diagnosis represented in our sample of surgery patients, we calculated the rate of admission (i.e., admissions/population) in each HRR. We then calculated the between-HRR variation for each ICD-9 admission rate and divided the sample into three groups: low-, medium-, and high-variation surgeries. Examples of low-variation procedures are hip and hernia surgery, and examples of high-variation procedures are skin graft and spinal fusion. We then obtained separate instrumental variables estimates for these three groups of surgery patients. Our estimates were very similar across groups and thereby provided evidence inconsistent with a significant selection problem in which EOL measures are correlated with patient selection. If there were significant selection, we would expect the estimates to differ across these three types of surgical patients. The reason is that some previous research indicated that aggressive hospitals, as measured by EOL measures, treat a greater proportion of the population. If so, these (patient) compositional differences would tend to affect estimates and result in different estimates across the three types of patients. Our evidence is inconsistent with such a selection story.

Data

Our data for our analysis comes primarily from the Medicare Provider Analysis and Review File (MEDPAR), which contains information on principal and secondary diagnoses, age, sex, comorbidities, and discharge status, including dates of death. The sample includes all Medicare patients admitted to short-term, acute-care, general U.S. nonfederal hospitals between July 1, 2000, and June 30, 2005, with a principal diagnosis of AMI, CHF, gastrointestinal bleeding, or stroke or with a diagnosis-related group classification of general, orthopedic, or vascular surgery.[11] We excluded patients from hospitals that were not in continuous operation during the period (9,600 patients from thirty-three hospitals) and from small hospitals—those with fewer than 350 admissions in any year (40,756 patients from 1,640 hospitals). We also excluded patients younger than sixty-six years ($n = 1{,}562{,}532$) because we used the period 180 days before admission to construct the measures of patient health that we used as risk adjusters (X variables in the preceding two equations) in the analysis. Patients older than ninety years ($n = 720{,}191$) were excluded as well because salvage treatment may not be given with the same intensity in this age group. Patients enrolled in managed care plans were excluded because there are no claims data for these patients. Finally, we excluded patients whose dates of death preceded their discharge dates ($n = 488$), who were transferred in from other hospitals ($n = 13{,}615$), and AMI or stroke patients discharged alive in fewer than two days ($n = 50{,}926$). The last exclusion was based on evidence that such early discharge may not represent actual AMIs or strokes (Romano, Remy, and Luft 1996).

We used only the first admission for each patient during this period. The first admission is defined as one without a prior admission for the medical condition or surgical category in the last five years (using data from July 1, 1995, or for patients younger than seventy back to when the person turned sixty-five and entered the Medicare system). Based on all these selection criteria, our sample consisted of 8,529,595 patients from 3,321 hospitals.

The dependent variable for both the medical and surgical patients was mortality within thirty days of admission. We limited the surgical sample to patients who experienced an in-hospital complication. Complications are defined as events noted in the discharge record that were likely to have developed after admission and were not present before the hospitalization. The forty-two complication categories ranged from cardiac arrest to wound infection, all of which increased the probability of death. A detailed description of these complications and a description of the construction of the complication list were provided by Silber and colleagues in 2007.[12] The medical conditions we selected were a subset of the Agency for Healthcare Research and Quality (AHRQ) Quality Indicators for which mortality was a relevant outcome. The sample sizes for patients with these specific illnesses were 776,773 for general surgery patients (with complication); 250,477 for vascular surgery patients (with complication); 912,500 for orthopedic patients (with complication); 1,127,861 for CHF patients; 931,551 for AMI patients; 891,118 for stroke patients; and 724,061 for patients admitted for GI bleeding.

The key explanatory variable is the total inpatient charges (spending associated with all resource use and tallied by hospital) for the Medicare admission, which come from the MEDPAR data. The instruments used to predict inpatient spending are three 2008 *Dartmouth Atlas of Health Care* (Wennberg et al. 2008) measures of each hospital's intensity of resource use during the last two years of life: days spent in the

ICU, days spent in hospital but not in the ICU, and total inpatient spending. All three *Dartmouth* end-of-life measures are closely correlated with actual inpatient spending associated with the admission of the patients in our sample. The *Dartmouth* end-of-life variables were constructed using a sample of all decedents with nine chronic illnesses: malignant cancer/leukemia, congestive heart failure, chronic pulmonary disease, dementia, diabetes with end organ damage, peripheral vascular disease, chronic renal failure, severe chronic liver disease, and coronary artery disease. Decedents were assigned to the hospital to which they were admitted most often, and the EOL measures were calculated for 2001 to 2005. All EOL measures were adjusted for differences in the patients' demographic characteristics (e.g., age and sex). Other explanatory variables included an extensive set of patient-level risk adjusters. These variables helped diminish the severity of reverse causality by providing a better characterization of the patients' health. We adopted the approach developed by Elixhauser and colleagues (1998), which is based on identifying twenty-seven comorbidities and has been shown to be highly predictive of mortality (Southern, Quan, and Ghali 2004; Stukenborg, Wagner, and Connors 2001). For surgical patients, we also adjusted for diagnosis-related groups (DRGs) and categories within DRGs that had homogenous risks. Comorbidities were constructed from data during the index hospitalization and on prior hospitalizations in the six-month period before admission.

Hospital characteristics also were included in the regression analysis. We used the resident-to-bed ratio, defined as the number of interns plus residents divided by the mean number of operational beds, as a measure of teaching intensity. The resident-to-bed ratio was obtained from the Medicare Cost Reports from CMS (HCRIS). Other hospital characteristics were technology level, hospital size, nurse-to-bed ratio, and share of registered nurses. These data come from the 2003 American Hospital Association (AHA) survey of hospitals. HCRIS data were also used to identify hospitals that merged, opened, or closed during the study period. Finally, we used the MEDPAR data to calculate the average total charges by hospital and year for surgery patients who did not experience a complication (were not in our outcome analysis). We calculated the average total charges separately for general surgery patients, vascular surgery patients, and orthopedic surgery patients. These average spending measures are intended to control for unmeasured hospital and patient characteristics that may be associated with mortality.

Estimates of Effects of Inpatient Spending and End-of-Life Measures on Thirty-Day Mortality

Table 1 shows estimates of the association between inpatient spending and thirty-day mortality for patients who were admitted to the hospital for general surgery, vascular surgery, and orthopedic surgery and who subsequently experienced a complication during the admission. Each estimate comes from a separate regression model, and all the estimates were obtained using several model specifications and estimation procedures. The estimates in column 1 are from a basic model that controls for only the year and subcategory of admission. In column 2, patient-level controls for medical risk factors are added to the model, and in column 3, hospital-level factors (e.g., resident-to-bed ratio) are added to the model. Estimates in columns 4 through 7 were obtained using a complete set of controls (patient and hospital) and instrumental variables (IV). Column 4's estimates were based on a model using decedents' ICU days and non-ICU hospital days as instruments for inpatient spending. The estimates in column 5 are from a model using only decedents' ICU days as an instrument, including non-ICU days as a covariate (just identified model). The estimates in column 6 are from a model that used total inpatient spending on decedents as an instrument, excluding ICU days and non-ICU hospital days as covariates. Finally, in column 7, we estimated the just identified model using total inpatient spending on decedents as an instrument, including ICU days and non-ICU hospital days as covariates.

Beginning with the sample of patients admitted for general surgery, the estimates in columns 1 through 3 indicate that greater resource use, as measured by inpatient spending, is associated with higher rates of thirty-day mortality, which is a result consistent with the "flat-of-the-curve" argument. The estimate in column 1 indicates that a $10,000 (20%) increase in spending is associated with a 0.29 percentage point (3%) increase in thirty-day mortality. Although adding patient- and hospital-level characteristics to the model (columns 2 and 3) reduces this positive association, the estimates remain small and statistically significant.[13] Columns 4 through 7 are instrumental variables estimates of the effect of inpatient spending on thirty-day mortality. To obtain these estimates, we used the *Dartmouth* EOL measures as instruments for actual inpatient spending. The first-stage, partial F-statistics are

TABLE 1
Estimates of the Association between Inpatient Spending ($10,000s) and Thirty-Day Morality Surgery Patients Experiencing a Complication

	(1)-OLS	(2)-OLS	(3)-OLS	(4) – IV	(5) – IV	(6) – IV	(7) – IV
General surgery patients	0.0029*	0.0010*	0.0016*	−0.0115*	−0.0088*	−0.0148*	−0.0249*
(mean of dep. var. = 0.11)	(0.0002)	(0.0001)	(0.0001)	(0.0023)	(0.0030)	(0.0024)	(0.0072)
Vascular surgery patients	0.0004**	−0.0011*	−0.0009*	−0.0145*	−0.0113*	−0.0132*	−0.0327*
(mean of dep. var. = 0.19)	(0.0002)	(0.0002)	(0.0002)	(0.0029)	(0.0028)	(0.0022)	(0.0073)
Orthopedic surgery patients	0.0045*	0.0024*	0.0037*	−0.0168*	−0.0106*	−0.0208*	−0.0069*
(mean of dep. var. = 0.06)	(0.0003)	(0.0002)	(0.0002)	(0.0023)	(0.0040)	(0.0023)	(0.0022)
End-of-life instruments:							
ICU days				Excluded Instrument	Excluded Instrument		Included Covariate
Non-ICU hospital days				Excluded Instrument	Included Covariate		Included Covariate

Continued

TABLE 1—*Continued*

	(1)-OLS	(2)-OLS	(3)-OLS	(4) – IV	(5) – IV	(6) – IV	(7) – IV
Inpatient spending (Dartmouth)						Excluded Instrument	Excluded Instrument
Partial F-excluded instruments							
General surgery patients				14*	28*	37*	27*
Vascular surgery patients				18*	26*	21*	29*
Orthopedic surgery patients				14*	27*	35*	53*
Patient characteristics	No	Yes	Yes	Yes	Yes	Yes	Yes
Hospital characteristics	No	No	Yes	Yes	Yes	Yes	Yes

Notes: Values in each cell are an estimate from a separate regression.
1. Estimates in column 1 are from a regression model that includes year dummy variables and dummy variables for admission type (e.g., separate dummy variables for subcategories of admission within general surgery category).
2. Estimates in column 2 add patient-level characteristics (i.e., patient risk adjustment) to regression model.
3. Estimates in column 3 add to regression model hospital-level characteristics (e.g., teaching hospital and average total spending on surgery patients without a complication).
4. Estimates in columns 4 through 7 are from instrumental variables procedure. In these regressions, we replace the actual inpatient spending with its predicted value. Instruments used to predict inpatient spending are end-of-life measures and are indicated in table.
5. Robust (Huber sandwich estimator) errors that allow for nonindependence of observations within a hospital are in parentheses.
6. $^*p < .01$; $^{**}p < .05$.

shown in the table and indicate a relatively strong first-stage correlation. The IV estimates indicate that greater inpatient spending is associated with lower incidence of thirty-day mortality of general surgery patients. The estimate in column 4 indicates that a \$10,000 (20%) increase in inpatient spending is associated with a 1.15 percentage point (11%) decrease in thirty-day mortality. The estimate in column 5 implies a smaller effect: a \$10,000 increase in inpatient spending is associated with a 0.88 percentage point (8%) decrease in thirty-day mortality. Finally, the estimate in column 7 implies the largest effect: a \$10,000 increase in inpatient spending is associated with a 2.49 percentage point (24%) decrease in thirty-day mortality. These are nontrivial effect sizes and inconsistent with the "flat-of-the-curve" hypothesis.

The next panel of table 1 reports estimates for the sample of patients admitted for vascular surgery. For these patients, the estimates in column 1 indicate that greater inpatient spending is positively associated with thirty-day mortality but that the effect size is minuscule. Adding patient and hospital controls reverses the sign of the association between spending and thirty-day mortality, but the estimates remain small in magnitude.[14] The IV estimates of the effect of inpatient spending using the EOL measures as instruments indicate that greater spending is associated with lower thirty-day mortality. A \$10,000 (17%) increase in inpatient spending is associated with between a 1.1 (column 5) and a 3.3 (column 7) percentage point decrease in thirty-day mortality. In relative terms, these estimates indicate that a 17 percent increase in inpatient spending is associated with a 6 to 18 percent decrease in thirty-day mortality. Clearly, these estimates reject the "flat-of-the-curve" hypothesis for this outcome.

The last panel of table 1 presents results for patients admitted for orthopedic surgery. In this case, we discuss only the IV estimates because OLS estimates indicated similar conclusions for general and vascular surgery patients. The IV estimates indicate that a \$10,000 (approximately 30%) increase in inpatient spending is associated with between a 0.7 (column 7) and a 2.1 (column 6) percentage point decrease in thirty-day mortality. Relative to the mean rate of thirty-day mortality, these are large effect sizes: between 11 and 33 percent.

We now turn to estimates of the association between inpatient spending and the thirty-day mortality of patients admitted for medical conditions: congestive heart failure (CHF), acute myocardial infarction (AMI), stroke, and gastrointestinal bleeding (GI bleeding). Starting with the

CHF patients (top panel), the estimates in columns 1 through 3 of table 2 indicate that greater inpatient spending is associated with the higher thirty-day mortality of CHF patients.[15] The IV estimates show that a $10,000 (50%) increase in inpatient spending is associated with between a 2.1 percentage point (25%) and a 3.0 percentage point (38%) decrease in the thirty-day mortality of CHF patients.

The next panel of table 2 shows the estimates of the association between spending and AMI. Here we found a markedly different pattern of estimates. The estimates in columns 1 through 3 suggest that greater inpatient spending is associated with lower thirty-day mortality, although the estimates are small. Two of four IV estimates indicate that greater inpatient spending is not significantly associated with mortality, results that are consistent with the "flat-of-the-curve" hypothesis.[16] But the IV estimates in columns 6 and 7 indicate that $10,000 of additional spending is associated with a 0.9 and 1.3 percentage point (6 to 9 percent) decrease in thirty-day mortality. It is clear that the results related to AMI are different from those of all other outcomes. Why this is the case is uncertain, as the data in table A2 of the appendix do not suggest that the IV procedure is any less valid for AMI than for other outcomes, although the partial F-statistics associated with the instruments are the smallest for AMI versus other types of patients.

The next two panels of table 2 show the results pertaining to mortality for patients admitted for stroke and GI bleeding. The estimates in column 1 indicate that inpatient spending is positively related to stroke patients' mortality. But when we add controls for patient severity and hospital characteristics, this outcome is reversed, although the magnitude of the estimate remains small.[17] The IV estimates suggest that a $10,000 (50%) increase in inpatient spending is associated with between a 4.0 percentage point (24%) and a 5.5 percentage point (31%) decrease in mortality from stroke. For GI bleeding, the findings are much the same as for most other outcomes. In OLS models, inpatient spending is positively associated with mortality of these patients. Controlling for patient severity only partly eliminates this problem. The IV estimates indicate a beneficial effect of spending: a $10,000 (60%) increase in inpatient spending is associated with between a 1.1 percentage point (18%) and a 2.7 percentage point (44%) decrease in the thirty-day mortality of patients admitted for GI bleeding.

Up to this point, we assumed that the effects of spending on mortality were uniform across patient types and hospitals. But additional

TABLE 2
Estimates of the Association between Inpatient Spending ($10,000s) and Thirty-Day Mortality of Medical Patients

	(1)-OLS	(2)-OLS	(3)-OLS	(4) – IV	(5) – IV	(6) – IV	(7) – IV
CHF patients	0.0034*	0.0025*	0.0032*	−0.0212*	−0.0255*	−0.0222*	−0.0302*
(mean of dep. var. = 0.08)	(0.0002)	(0.0002)	(0.0002)	(0.0037)	(0.0042)	(0.0034)	(0.0087)
AMI patients	−0.0016*	−0.0019*	−0.0022*	−0.0003	0.0084	−0.0086*	−0.0132*
(mean of dep. var. = 0.15)	(0.0001)	(0.0001)	(0.0001)	(0.0045)	(0.0055)	(0.0028)	(0.0043)
Stroke patients	0.0018*	−0.0001	−0.0004**	−0.0549*	−0.0468*	−0.0516*	−0.0399*
(mean of dep. var. = 0.17)	(0.0002)	(0.0002)	(0.0002)	(0.0084)	(0.0109)	(0.0072)	(0.0157)
GI-bleeding patients	0.0076*	0.0050*	0.0057*	−0.0116*	−0.0117*	−0.0137*	−0.0274*
(mean of dep. var. = 0.06)	(0.0004)	(0.0003)	(0.0003)	(0.0026)	(0.0042)	(0.0027)	(0.0087)
End-of-life instruments:							
ICU days				Excluded Instrument	Excluded Instrument		Included Covariate
Non-ICU hospital days				Excluded Instrument	Included Covariate		Included Covariate
Inpatient spending (Dartmouth)						Excluded Instrument	Excluded Instrument
Partial F-excluded instruments							
CHF patients				20*	35*	66*	21*
AMI patients				13*	15*	35*	23*
Stroke patients				21*	32*	73*	13*
GI-bleeding patients				26*	33*	96*	24*
Patient characteristics	No	Yes	Yes	Yes	Yes	Yes	Yes
Hospital characteristics	No	No	Yes	Yes	Yes	Yes	Yes

Notes: See notes to table 1. In addition, for medical patients, we include among the hospital characteristics separate measures of the average total spending on three types of surgery patients who did not experience a complication: general, vascular, and orthopedic.

spending may be more or less effective for patients of varying severity or across different types of hospitals. For example, Chandra and Staiger (2007) demonstrated that additional spending for treatment of AMI is much more effective for younger (<80) than for older patients. Here we examined whether the effect of spending differed by age (<80, ≥80) and whether the patient had a previous diagnosis of cancer, and also the hospital's resident-to-bed ratio (<0.25, ≥0.25), which is a measure of the hospital's teaching intensity. Table 3 shows the IV estimates (from column 5's specification of tables 1 and 2) of the effect of spending by patient and hospital characteristics.

In general, the estimates in table 3 reveal the near absence of heterogeneity in the effect of spending by patient type. For example, spending is not systematically less effective for older patients or for patients diagnosed with cancer. In only a couple of cases did the estimates of the effect of spending differ much by patient type. For vascular surgery and CHF, more spending appears to have a larger beneficial effect on those patients without cancer. In contrast, we see a systematic difference in the effects of spending by hospital type, as classified according to the resident-to-bed ratio, which is a measure of teaching intensity. The effects of spending are larger (more negative), in fact sometimes twice as large, in less teaching-intensive hospitals compared with more teaching-intensive hospitals. For example, an additional $10,000 in spending at nonteaching hospitals is associated with an 11 percent decrease in the thirty-day mortality of general surgery patients, a 17 percent decrease in the thirty-day mortality of orthopedic surgery patients, a 6 percent decrease in the thirty-day mortality of vascular surgery patients, and a 25 to 33 percent decrease in the thirty-day mortality of patients admitted for CHF, GI bleeding, or stroke.

Conclusions

For all patients except those admitted for AMI, the instrumental variables estimates consistently indicated that greater inpatient spending was associated with a lower thirty-day mortality for surgical patients who experienced an in-hospital complication as well as medical patients (CHF, stroke, GI bleeding). Table 4 summarizes our findings. Overall, the IV estimates (from column 5 of tables 1 and 2) indicated that a 10 percent increase in inpatient spending is associated with between a

TABLE 3
IV Estimates of the Association between Inpatient Spending ($10,000s) and Thirty-Day Mortality by Patient and Hospital Characteristics

	General Surgery	Orthopedic Surgery	Vascular Surgery	AMI	Stroke	GI-Bleeding	CHF
Patient's age							
Age < 80	−0.0082*	−0.0073	−0.0102**	0.0032	−0.0500**	−0.0087*	−0.0220**
	(0.0036)	(0.0045)	(0.0031)	(0.0057)	(0.0110)	(0.0041)	(0.0045)
Mean of dependent variable	0.0856	0.0385	0.1644	0.1101	0.1402	0.0484	0.0632
Age 80+	−0.0097**	−0.0124**	−0.0134**	0.0142*	−0.0424**	−0.0148*	−0.0287**
	(0.0034)	(0.0047)	(0.0039)	(0.0069)	(0.0121)	(0.0064)	(0.0047)
Mean of dependent variable	0.1476	0.0943	0.2375	0.2066	0.2160	0.0786	0.1075
Patient comorbidity							
Has cancer	−0.0112*	−0.0167*	−0.0042	−0.0115	−0.0393*	0.0179	−0.0152
	(0.0066)	(0.0095)	(0.0086)	(0.0133)	(0.0203)	(0.0187)	(0.0095)
Mean of dependent variable	0.1775	0.2095	0.2276	0.2605	0.2986	0.1936	0.1803
Does not have cancer	−0.0080**	−0.0095**	−0.0116**	0.0099*	−0.0471**	−0.0141**	−0.0261**
	(0.0031)	(0.0039)	(0.0028)	(0.0058)	(0.0107)	(0.0040)	(0.0044)
Mean of dependent variable	0.0948	0.0571	0.1859	0.1429	0.1687	0.0520	0.0785
Hospital type							
Resident-to-bed ratio> = 0.25	0.0007	−0.0085	−0.0071	0.0141	−0.0233	−0.0012	−0.0151**
	(0.0066)	(0.0060)	(0.0055)	(0.0144)	(0.0164)	(0.0076)	(0.0054)
Mean of dependent variable	0.1034	0.0570	0.1772	0.1404	0.1909	0.0666	0.0740
Resident-to-bed ratio < 0.25	−0.0116**	−0.0111**	−0.0124**	0.0077	−0.0514**	−0.0148**	−0.0288**
	(0.0033)	(0.0047)	(0.0032)	(0.0060)	(0.0101)	(0.0046)	(0.0053)
Mean of dependent variable	0.1099	0.0647	0.1913	0.1492	0.1722	0.0618	0.0859

Notes: See notes to table 1, except that $^{*}p < .05$; $^{**}p < .01$. In addition, IV estimates are from a model corresponding to column 5 in tables 1 and 2.

TABLE 4
Summary of Estimates of Associations between Inpatient Spending and Thirty-Day Mortality Associated with a 10 Percent Increase in Inpatient Spending

Patients	General Surgery	Vascular Surgery	Orthopedic Surgery	CHF	AMI	Stroke	GI Bleeding
Percent decrease in 30-day mortality associated with 10 percent increase in inpatient spending	4.0 to 11.3	3.6 to 10.4	3.7 to 11.1	5.1 to 7.2	0 to 3.0	4.6 to 6.3	3.1 to 7.3

3.1 and an 11.3 percent decrease in thirty-day mortality, depending on the type of patient. Consider the trade-off between spending and mortality implied by the estimates in table 4. These estimates suggest that elderly patients at the end of life can spend 10 percent more—in our data, between $2,000 and $5,000—in return for a 3.1 to an 11.3 percent increase in survival. We do not think it unreasonable for people to be willing to make this trade-off, although we do not know the exact cost of each year of life in these cases.

More important, our results suggest that even though it may be cost-effective to eliminate some portion of inpatient spending, this reduction would come at a considerable cost for survival, at least for inpatients. Except for AMI patients, greater spending on inpatients is associated with improved survival, particularly at the less teaching-intensive hospitals that are numerically dominant. Why the AMI results differ from other outcomes is unclear. One possible explanation is that the negative association between spending and the thirty-day mortality of AMI patients obtained from noninstrumental variable models may simply reflect the fact that the most effective AMI treatment includes the early use of invasive technology (e.g., see Kostis et al. 2007). If hospitals initially use invasive technology, they will spend more per patient and they will have better outcomes in AMI. This is not generally true for congestive heart failure (or other outcomes), which is a chronic problem that usually requires less expensive technology initially on admission but that may require time in the ICU if the initial treatment fails. The sicker CHF patients stay in the ICU, and they also are more likely to die. Therefore, the instrument is possibly more valid in this case than in AMI, whose typical patients need the benefit of intensive treatment, and that treatment is more prevalent in hospitals that spend more. In other words, in AMI, high spending is a signal that the patients received the more extensive early treatment, not necessarily that they were sicker. In CHF (and other outcomes), however, high spending is a signal that the patient's condition required more extensive treatment.

Based on our results, the narrowest interpretation of the "flat-of-the-curve" hypothesis—that health care spending could be reduced by 20 to 30 percent without adverse health effects—may be seriously misleading. In addition, such arguments may be misinterpreted as meaning that there are no benefits of greater spending. Instead, greater inpatient spending may be beneficial, and other types of spending (e.g., imaging, specialist visits), which may be more prone to supply-sensitive variation, may be particularly wasteful. Similarly, the popular view that more treatment is harmful, as expressed in the earlier quoted passage from *Consumer Reports*, may be grossly misleading. This conclusion is clearly not appropriate, at least according to our estimates and estimates from other recent studies (e.g., Chandra and Staiger 2007).

In the case of teaching hospitals, defined as having a resident-to-bed ratio greater than 0.25, we found that additional spending was less beneficial for survival than in nonteaching hospitals. This finding may reflect the generally higher levels of spending in more teaching-intensive hospitals, as well as their dual mission as centers for teaching, which increases costs (spending), and providers of patient care. The relative inefficiency of these hospitals may be a measure of the cost of training physicians. These findings also are consistent with the data presented in Lindenauer and colleagues (2007), demonstrating that teaching hospitals generally have a relatively high quality of patient care and that the quality of patient care is less responsive to financial incentives than it is in nonteaching hospitals.

In sum, we analyzed the association between inpatient spending—greater resource use—and thirty-day mortality. We tried to extend the literature by examining the effect of inpatient medical spending on the mortality of Medicare patients admitted to the hospital for surgery (general, orthopedic, vascular) and of patients admitted for medical conditions (AMI, CHF, stroke, and GI bleeding) between 2001 and 2005. Using a sample of surgical patients who experienced an in-hospital health shock is new in this context and partly addresses the reverse causality problem that is likely to bias estimates. We further addressed this issue by using an instrumental variables procedure using evidence showing that the intensity of treatment and resource use for patients in a hospital is strongly associated with the intensity of treatment for patients at the end of life (decedents). Even though this approach seemed reasonable based on the evidence we presented, hospital-specific factors may have been omitted that affect both inpatient spending and survival. Thus, the IV estimates have to be interpreted with appropriate caution.

Given this caveat, we found that except for AMI patients, greater inpatient spending is significantly and negatively associated with thirty-day mortality. The increase in survival was nontrivial: a 10 percent increase in inpatient spending is associated with a 3.1 to 11.3 percent increase in survival. These findings add to other recent studies showing that greater spending is beneficial (Card, Dobkin, and Maestas 2008; Chandra and

Staiger 2007; Doyle 2005, 2007). This growing body of literature raises questions about the general applicability of the "flat-of-the-curve" argument, at least for inpatient spending.[18] While the U.S. health care system undoubtedly has some inefficiency and waste, the amount may not be as large as commonly believed, at least for hospitalized Medicare patients. We argued that there appear to be market mechanisms that would work to eliminate such waste and that the incentive to do so grows with the amount of waste. Spending 20 to 30 percent more than is necessary, indeed without any benefit, seems to be sufficiently large for such incentives to spur action. But if the amount of waste is significantly less than 20 to 30 percent, for example, only 5 percent, then there is relatively little problem and no mystery as to why market mechanisms have not eliminated the problem—because the "flat-of-the-curve" argument may not be particularly accurate. Clearly, more research is needed, particularly research that provides a credible assessment of the causal relationship between spending and health. Even though casual observation suggests that there is inefficiency, such beliefs are inconsistent with economic incentives and a growing body of empirical evidence.

Endnotes

1. These figures come from projections for 2010 of the Centers for Medicare and Medicaid Services (CMS), available at http://www.cms.gov/NationalHealthExpendData/downloads/proj2009.pdf (accessed July 14, 2010).
2. For example, health care spending in large, self-insured firms tends to be more or less the same and grows just as rapidly as spending in government programs (Kaiser Family Foundation 2008). In addition, evidence suggests that managed care contracting has relatively small effects on spending in both private and public settings (e.g., Chernew, DeCicca, and Town 2008; Fisher et al. 2009; Glied 2000).
3. The heterogeneous effects of spending with respect to geographic area (more- or less-intensive treatment areas) and suitability of patients for treatment reported by Chandra and Staiger (2007) are consistent with the "flat-of-the-curve" hypothesis. However, on average, the estimates by Chandra and Staiger (2007) indicate that the spending is very effective: approximately $9,000 per life year saved. Doyle (2005, 2007) reaches a similar conclusion in his studies of the associations between inpatient spending and mortality. In their study of the effect of obtaining Medicare on the general population's mortality, Card, Dobkin, and Maestas (2008) also report results that imply a relatively low cost per year of life saved.
4. We used a thirty-day mortality, as opposed to longer periods, because it is the measure most closely matched to our research question. The use of longer periods, for example, a one-year mortality, risks confounding the effects of inpatient spending in the treating hospital with the effects of spending and treatment by other providers.
5. We found that logistic regression methods produced very similar estimates of the effect of a change in inpatient spending on the probability of thirty-day mortality (i.e., the marginal

effect on the probability). We prefer ordinary least squares because it facilitates the use of the instrumental variables procedure.

6. We also obtained estimates of associations between end-of-life measures and 60, 90, and 365-day mortality for surgical patients (reduced form analyses of surgical patients). We have reported these results elsewhere (Silber et al. 2010). The key findings from these analyses are that end-of-life measures of resource use remain negatively associated with mortality but that associations between end-of-life measures decline as the window measuring mortality opens. This is exactly what one would expect if beyond thirty days, there were no differences in the mortality hazard between patients treated at hospitals with higher or lower end-of-life resource use. The equal post-thirty-day mortality rate across hospitals dilutes the initial advantage obtained by patients in hospitals that treat intensively (greater resource use, as measured by end-of-life measures).
7. We used average spending on surgery patients who do not have a complication while in hospital. We discuss this variable in more detail later.
8. Alternatively and more problematically for the validity of the IV strategy, the variation in EOL spending may be due to the type of specialization of providers and sorting of patients suggested by Chandra and Staiger (2007). We acknowledge this possibility and try to address it by controlling for detailed patient and hospital characteristics. Estimates are not sensitive to the inclusion of patient characteristics.
9. Silber, Rosenbaum and Ross (1995) have shown this formally using the omega statistic, and Silber has duplicated these findings in a paper using the same Medicare claims data set as described in this article (see Silber et al. 2007). The omega statistic describes the relative contribution of patient characteristics (severity) to hospital factors in predicting an outcome: A larger statistic indicates that patient factors are relatively more important. The omega statistic associated with predicting the probability that a surgery patient will experience a complication is approximately 300, and the omega statistic associated with predicting the thirty-day mortality of all surgery patients is also about 300. In contrast, the omega statistic associated with predicting the thirty-day mortality of surgery patients experiencing a complication is 90 (the difference in omegas was statistically significant). In other words, the thirty-day mortality of surgery patients experiencing a complication is less influenced by severity than by the thirty-day mortality or the probability of experiencing a complication of (all) surgery patients.
10. The regression models included a full set of covariates.
11. This section draws heavily from Volpp and colleagues (2007). The data used in our study are the same as that used in this article. We excluded patients in hospitals that did not have Medicare Cost Report data (14,514 patients from 39 hospitals), that were missing more than two months of data at various times during the period (276,040 patients from 703 hospitals), and whose hospitalization included July 1, 2003 (n = 26,856). These exclusions were related to the objective in the study by Volpp and colleagues (2007), which evaluated the effects of the resident hour rule change that took effect in July 2003 and should not affect the current analysis.
12. We included in the sample those surgical patients who died but for whom the discharge record did not indicate a complication. Fewer than 5 percent of deaths were not documented to have had a preceding complication (Silber et al. 2007).
13. In contrast, estimates of the association between the EOL measures and thirty-day mortality (see tables A3 and A4) always are negative and statistically significant, which differs fundamentally from the results of previous studies that have conducted similar analyses (Fisher et al. 2003; Skinner, Fisher, and Wennberg 2005). The estimates in column 3 of table A3 indicate that the three additional ICU days for decedents, which represent a change from the 25th to 75th percentile in the distribution, is associated with a 0.24 percentage point (2%) decrease in the thirty-day mortality.

14. The estimates associated with the EOL measures are negatively and significantly associated with thirty-day mortality (tables A3 and A4), and the addition of patient and hospital covariates has little effect on estimates, which bolsters the case for the validity of the IV procedure. In column 3 of table A3, the estimate associated with ICU days indicates that the three additional ICU days are associated with a 0.78 percentage point (4%) decrease in the thirty-day mortality.
15. All the estimates of the association between the EOL measures and thirty-day mortality are negative and statistically significant (tables A3 and A4). The estimate in column 3 of table A3 indicates that three additional ICU days are associated with a 0.57 percentage point (6%) decrease in the thirty-day mortality of CHF patients. Here, too, we observed that the addition of covariates has little effect on the EOL estimates, which support the IV procedure.
16. Some of the estimated associations between end-of-life measures and 30-day mortality also are positive (see tables A3 and A4).
17. The EOL measures are always negatively and significantly associated with the mortality of stroke patients. In this case, the Dartmouth estimates are somewhat affected by the addition of hospital controls, but not patient controls.
18. Note that even if we constructed a weighted average of the effects listed in table 4, including zero for AMI, we would still find a substantial effect of inpatient spending on mortality. AMI patients represented approximately one-sixth of all patients in our analysis.

References

Card, D., C. Dobkin, and N. Maestas. 2008. The Impact of Nearly Universal Insurance Coverage on Health Care Utilization: Evidence from Medicare. *American Economic Review* 98(5):2242–58.

Chandra, A., and D. Staiger. 2007. Productivity Spillovers in Health Care: Evidence from the Treatment of Heart Attacks. *Journal of Political Economy* 115(1):103–40.

Chernew, M., P. DeCicca, and R. Town. 2008. Managed Care and Medical Expenditures of Medicare Beneficiaries. *Journal of Health Economics* 27(6):1451–61.

Consumer Reports Health.org. 2008 (July). Too Much Treatment? Aggressive Medical Care Can Lead to More Pain with No Gain. Available at http://cdn.consumerreports.org/health/doctors-hospitals/hospital-care/aggressive-medical-care/overview/aggressive-medical-care-ov.htm (accessed October 7, 2010).

Dartmouth Atlas of Health Care. 2008. Health Care Spending, Quality and Outcomes. Available at http://www.dartmouthatlas.org/downloads/reports/Spending_Brief_022709.pdf (accessed October 7, 2010).

Doyle, J. 2005. Health Insurance, Treatment and Outcomes: Using Auto Accidents as Health Shocks. *Review of Economics and Statistics* 87(2):256–70.

Doyle, J. 2007. Returns to Local-Area Health Care Spending: Using Health Shocks to Patients Far from Home. NBER working paper, number 13301. National Bureau of Economic Research: Cambridge, MA. Available at http://www.nber.org/papers/w13301 (accessed October 13, 2010).

Elixhauser, A., C. Steiner, D. Harris, and R. Coffey. 1998. Comorbidity Measures for Use with Administrative Data. *Medical Care* 36(1):8–27.

Fisher, E., D. Goodman, J. Skinner, and K. Bronner. 2009. Health-care Spending, Quality, and Outcomes. More Isn't Always Better. Available at http://www.dartmouthatlas.org/downloads/reports/Spending_Brief_022709.pdf (accessed July 19, 2010).

Fisher, E., J. Wennberg, T. Stukel, D. Gottlieb, F. Lucas, and E. Pinder. 2003. The Implications of Regional Variations in Medicare Spending. Part 2: Health Outcomes and Satisfaction with Care. *Annals of Internal Medicine* 138(4):288–98.

Fuchs, V.R. 1966. The Contribution of Health Services to the American Economy. *The Milbank Memorial Fund Quarterly* 44(4):65–103, Part 2.

Fuchs, V. 2004. More Variation in Use of Care, More Flat-of-the-Curve Medicine. *Health Affairs* 10(1377):104–7.

Glied, S. 2000. Managed Care. In *Handbook of Health Economics*, ed. J. Newhouse and A. Culyer, 708–45. Amsterdam: Elsevier.

Gruber, J. 2006. The Role of Consumer Copayments for Health Care: Lessons from the RAND Health Insurance Experiment and Beyond. Washington, DC: Kaiser Family Foundation. http://www.kff.org/insurance/7566.cfm (accessed October 7, 2010).

Kaiser Family Foundation. 2008. Health Care Costs and Election. Available at http://www.kff.org/insurance/h08_7828.cfm (accessed October 8, 2010).

Kostis, W.J., K. Demissie, S.W. Marcella, Y.-H. Shao, A.C. Wilson, and A.E. Moreyra. 2007. Weekend versus Weekday Admission and Mortality from Myocardial Infarction. *New England Journal of Medicine* 356:1099–1109.

Lindenauer, P.K., D. Remus, S. Roman, M.B. Rothberg, E.M. Benjamin, and D.W. Bratzler. 2007. Public Reporting and Pay for Performance in Hospital Quality Improvement. *New England Journal of Medicine* 356:486–96.

Newhouse, J. 1993. *Free for All*. Cambridge, MA: Harvard University Press.

Romano, P.S., L.L. Remy, and H.S. Luft. 1996. Second Report of the California Hospital Outcomes Project (1996): Acute Myocardial Infarction Volume Two: Technical Appendix (table of content).

Available at http://escholarship.org/uc/chsrpc_coshpd#1996 (accessed October 8, 2010).

Silber, J.H., R. Kaestner, O. Even-Shoshan, Y. Wang, and L. Bressler. 2010. Aggressive Treatment Style and Surgical Outcomes. *Health Services Research* 45:1872–92.

Silber, J.H., P. Romano, A. Rosen, Y. Wang, O. Even-Shoshan, and K. Volpp. 2007. Failure-to-Rescue: Comparing Definitions to Measure Quality of Care. *Medical Care* 45(10):918–25.

Silber, J.H., P.R. Rosenbaum, and R. Ross. 1995. Comparing the Contributions of Groups of Predictors: Which Outcomes Vary with Hospital Rather than Patient Characteristics? *Journal of the American Statistical Association* 90:7–18.

Silber, J.H., P.R. Rosenbaum, M.E. Trudeau, W. Chen, X. Zhang, R.R. Kelz, R.E. Mosher, and O. Even-Shoshan. 2005. Changes in Prognosis after the First Postoperative Complication. *Medical Care* 43(2):122–31.

Silber, J.H., S. Williams, H. Krakauer, and S. Schwartz. 1992. Hospital and Patient Characteristics Associated with Death after Surgery: A Study of Adverse Occurrence and Failure to Rescue. *Medical Care* 30(7):615–29.

Skinner, J., E. Fisher, and J. Wennberg. 2005. The Efficiency of Medicare. In *Analyses in the Economics of Aging*, ed. D.A. Wise, 129–60. Chicago: University of Chicago Press.

Skinner, J., and D. Staiger. 2009. Technology Diffusion and Productivity Growth in Health Care. NBER Working Paper 14865. Cambridge, MA: NBER.

Southern, D., H. Quan, and W. Ghali. 2004. Comparison of the Elixhauser and Charlson/Deyo Methods of Comorbidity Measurement in Administrative Data. *Medical Care* 42(4):355–60.

Stukenborg, G., D. Wagner, and A. Connors. 2001. Comparison of the Performance of Two Comorbidity Measures, With and Without Information from Prior Hospitalizations. *Medical Care* 39(7):727–39.

Volpp, K., A. Rosen, P.R. Rosenbaum, P. Romano, O. Even-Shoshan, Y. Wang, L. Bellini, T. Behringer, and J.H. Silber. 2007. Mortality among Hospitalized Medicare Beneficiaries in the First 2 Years Following ACGME Resident Duty Hour Reform. *Journal of the American Medical Association* 298(9):975–83.

Wennberg, J., E. Fisher, D. Goodman, and J. Skinner. 2008. Tracking the Care of Patients with Severe Chronic Illness. In *Dartmouth Atlas of Health Care*. Available at http://www.dartmouthatlas.org/downloads/atlases/2008_Chronic_Care_Atlas.pdf (accessed October 8, 2010).

Methodological Appendix

As noted in the text, to address the likely reverse causality between thirty-day mortality and inpatient spending, we used an instrumental variables approach. The validity of the instrumental variables procedure depends on two criteria: (1) that there is a variable, or set of variables, correlated with the endogenous variable, which in this case is inpatient spending, and (2) that this variable, or set of variables, does not belong in the regression model of interest, which in this case is the thirty-day mortality regression. By "belong in the regression," we mean that the variable, or set of variables, will have no effect (zero coefficient) on the outcome, given the other variables included in the regression. So in our case, hospital-specific measures of end-of-life resource use will have no effect on thirty-day mortality, given that we include inpatient spending, patient characteristics, and hospital characteristics in the regression model.

To determine whether our instruments met these criteria, we offered differences in the sample mean characteristics by whether a patient was in a hospital that was above or below the median value of two of the three end-of-life measures we used as instruments: ICU days and total hospital days (ICU and non-ICU days). A striking finding evident in tables A1 and A2 is that inpatient spending on patients (general surgery in table A1 and AMI in table A2) is strongly associated with both end-of-life measures and that inpatient spending on a general surgery patient is 20 to 30 percent of a standard deviation greater in hospitals that are above the median value than in those below the median value of these measures. We found this strong correlation between end-of-life measures and inpatient spending for all other patient types we analyzed. These findings are evidence that our instruments satisfied the first criterion. Tables A1 and A2 show the results of formal tests of this correlation.

Notably, the end-of-life measures were not significantly correlated with characteristics of patients' illness severity, including predicted mortality (summary of patient severity). Differences in patients' illness severity between patients in hospitals above or below the median end-of-life value are always less than or equal to 7 percent of a standard deviation (or of the mean in the case of dichotomous variables). In contrast, total inpatient spending on surgery patients who do not experience a complication (i.e., routine patients) is strongly correlated with

TABLE A1
Differences in Measured Characteristics by End-of-Life Measures: Sample of Patients Admitted for General Surgery Who Experienced a Complication

	Total (ICU and Non-ICU) Hospital Days			ICU Days		
	Above Median	Below Median	Difference/ Std. Dev.	Above Median	Below Median	Difference/ Std. Dev.
Inpatient spending (charges)	5.524	4.122	0.250	5.744	3.841	0.343
Thirty-day mortality	0.107	0.111	−0.013	0.108	0.110	−0.004
Patient characteristics						
Age	77.04	77.03	0.002	76.00	77.08	−0.013
Male	0.429	0.438	−0.018	0.433	0.432	0.002
Number of comorbidities	2.800	2.659	0.078	2.780	2.684	−0.053
Hypertension	0.542	0.516	0.053	0.537	0.523	0.029
COPD	0.232	0.236	0.008	0.234	0.234	−0.001
Diabetes	0.189	0.173	0.043	0.185	0.178	0.020
CHF	0.248	0.245	0.005	0.245	0.249	−0.011
Peripheral vascular disease	0.077	0.070	0.026	0.075	0.072	0.011
Renal failure	0.056	0.048	0.038	0.055	0.048	0.031
Predicted mortality	0.080	0.077	0.039	0.080	0.078	0.036
Hospital spending on patients not in analysis sample						
Average total charges general surgery patient	2.277	1.814	0.451	2.374	1.689	0.694
Average total charges vascular surgery patient	2.612	2.219	0.355	2.741	2.051	0.651
Average total charges orthopedic surgery patient	3.664	2.983	0.341	3.847	2.724	0.582

Notes: Difference/Std. Dev is the difference divided by the standard deviation if the variable is continuous. If it is a dichotomous variable, then this column reports the difference divided by the sample mean.
1. Average total charges for surgery patients are for patients without a complication and who are not part of the analyses.

TABLE A2
Differences in Measured Characteristics by End-of-Life Measures: Sample of Patients Admitted for AMI

	Total (ICU and Non-ICU) Hospital Days			ICU Days		
	Above Median	Below Median	Difference/ Std. Dev.	Above Median	Below Median	Difference/ Std. Dev.
Inpatient spending	4.336	3.263	0.216	4.544	2.974	0.320
Thirty-day mortality	0.151	0.148	0.009	0.150	0.150	0.001
Patient characteristics						
Age	77.34	77.25	0.023	77.31	77.35	−0.007
Male	0.507	0.528	−0.041	0.516	0.518	−0.004
Number of comorbidities	2.212	2.081	0.084	2.184	2.114	0.044
Hypertension	0.591	0.573	0.038	0.587	0.578	0.018
COPD	0.242	0.236	0.015	0.241	0.237	0.009
Diabetes	0.257	0.241	0.036	0.251	0.249	0.005
CHF	0.074	0.060	0.057	0.070	0.065	0.022
Peripheral vascular disease	0.104	0.103	0.003	0.104	0.104	0.001
Renal failure	0.073	0.063	0.040	0.072	0.064	0.031
Predicted mortality	0.152	0.147	0.056	0.151	0.149	0.021
Hospital spending on other patients						
Average total charge general surgery patient	2.213	1.793	0.427	2.301	1.672	0.664
Average total charge vascular surgery patient	2.534	2.182	0.330	2.653	2.021	0.621
Average total charge orthopedic surgery patient	3.598	3.001	0.313	3.763	2.760	0.545

Notes: Difference/Std. Dev is the difference divided by the standard deviation if the variable is continuous. If it is a dichotomous variable, then this column reports the difference divided by the sample mean.
1. Average total charges for surgery patients are for patients without a complication and who are not part of the analyses.

TABLE A3
Estimates of the Association between End-of-Life Measures (ICU and Non-ICU Days) and Thirty-Day Mortality: Reduced Form Estimates for Instrumental Variables Models of Column 4 in Tables 1 and 2 (samples of patients are the same as those used in tables 1 and 2)

	(1)-OLS	(2)-OLS	(3)-OLS
General surgery patients with complication			
ICU days	−0.0011**	−0.0015**	−0.0008**
	(0.0003)	(0.0003)	(0.0003)
Non-ICU hospital days	−0.0011**	−0.0011**	−0.0008**
	(0.0002)	(0.0002)	(0.0002)
Vascular surgery patients with complication			
ICU days	−0.0029**	−0.0030**	−0.0026**
	(0.0007)	(0.0006)	(0.0007)
Non-ICU hospital days	−0.0019**	−0.0021**	−0.0021**
	(0.0005)	(0.0005)	(0.0005)
Orthopedic surgery patients with complication			
ICU days	−0.0004**	−0.0007**	−0.0006*
	(0.0002)	(0.0002)	(0.0003)
Non-ICU hospital days	−0.0008**	−0.0008**	−0.0008**
	(0.0002)	(0.0002)	(0.0002)
CHF patients			
ICU days	−0.0019**	−0.0021**	−0.0019**
	(0.0004)	(0.0004)	(0.0005)
Non-ICU hospital days	−0.0012**	−0.0012**	−0.0010**
	(0.0003)	(0.0003)	(0.0004)
AMI patients			
ICU days	0.0016**	0.0007*	0.0005
	(0.0004)	(0.0003)	(0.0003)
Non-ICU hospital days	0.0003	−0.0004	−0.0005**
	(0.0004)	(0.0002)	(0.0002)
Stroke patients			
ICU days	−0.0026**	−0.0026**	−0.0036**
	(0.0007)	(0.0007)	(0.0011)
Non-ICU hospital days	−0.0025**	−0.0028**	−0.0035**
	(0.0007)	(0.0007)	(0.0009)
GI-bleeding patients			
ICU days	−0.0004**	−0.0007**	−0.0006*
	(0.0002)	(0.0002)	(0.0003)
Non-ICU hospital days	−0.0002*	−0.0005**	−0.0006**
	(0.0001)	(0.0002)	(0.0002)

Continued

TABLE A3—*Continued*

	(1)-OLS	(2)-OLS	(3)-OLS
Patient characteristics	No	Yes	Yes
Hospital characteristics	No	No	Yes

Notes: Values in each cell are estimates from a separate regression. For example, the estimates of associations in column 1 of first row are between ICU days and non-ICU days, and thirty-day mortality of patients admitted for general surgery and who experienced a complication.
1. Estimates in column 1 are from a regression model that includes year dummy variables and dummy variables for admission type (e.g., separate dummy variables for subcategories of admission within the general surgery category).
2. Estimates in column 2 add patient characteristics (i.e., patient risk adjustment) to the regression model.
3. Estimates in column 3 add hospital characteristics (e.g., teaching hospital and average total spending on surgery patients without a complication) to the regression model.
4. Robust standard errors that allow for nonindependence of observations within hospital are in parentheses.
5. $^{*}p < .01$; $^{**}p < .05$.

the end-of-life measures. These differences in inpatient spending may reflect differences in the "prices" that hospitals use to calculate "charges," differences in patient types, and/or differences in hospital characteristics that are otherwise unmeasured. Including these spending measures for patients not in the sample controls for these unmeasured differences.

Finally, tables A3 and A4 give the reduced form estimates of the associations between end-of-life measures and patients' thirty-day mortality. These estimates are an important part of the instrumental variables approach because the instrumental variables estimates of the association between inpatient spending and thirty-day mortality are (in the case of one instrument) equal to the ratio of the association between end-of-life spending and thirty-day mortality to the association between end-of-life inpatient spending and inpatient spending. Thus, the sign of the association between end-of-life spending and thirty-day mortality reveals the sign of the instrumental variables estimate of the association between inpatient spending and thirty-day mortality. We also can determine whether the reduced form estimates of associations between end-of-life measures and patients' thirty-day mortality are sensitive to the addition of other covariates.

A striking finding revealed by the estimates in tables A3 and A4 is that all the estimates are negative and statistically significant except for those associated with AMI patients. More important, estimates are not very sensitive to the addition of patient or hospital characteristics.

TABLE A4
Estimates of the Association between End-of-Life Measures (inpatient spending) and Thirty-Day Mortality: Reduced Form Estimates for Instrumental Variables Models of Column 6 in Tables 1 and 2 (samples of patients are the same as those used in tables 1 and 2)

	(1)-OLS	(2)-OLS	(3)-OLS
General surgery patients with complication			
Inpatient spending (Dartmouth)	−0.0069*	−0.0072*	−0.0048*
	(0.0009)	(0.0009)	(0.0008)
Vascular surgery patients with complication			
Inpatient spending (Dartmouth)	−0.0089*	−0.0094*	−0.0099*
	(0.0011)	(0.0011)	(0.0016)
Orthopedic surgery patients with complication			
Inpatient spending (Dartmouth)	−0.0036*	−0.0042*	−0.0045*
	(0.0005)	(0.0006)	(0.0010)
CHF patients			
Inpatient spending (Dartmouth)	−0.0064*	0.0065*	−0.0062*
	(0.0010)	(0.0009)	(0.0014)
AMI patients			
Inpatient spending (Dartmouth)	0.0014	−0.0013**	−0.0023*
	(0.0009)	(0.0006)	(0.0007)
Stroke patients			
Inpatient spending (Dartmouth)	−0.0053*	−0.0060*	−0.0148*
	(0.0012)	(0.0012)	(0.0031)
GI-bleeding patients			
Inpatient spending (Dartmouth)	−0.0003	−0.0022*	−0.0031*
	(0.0004)	(0.0004)	(0.0008)
Patient characteristics	No	Yes	Yes
Hospital characteristics	No	No	Yes

Notes: See notes to table A3.

These two findings support the plausibility of the instrumental variables procedure.

In sum, the estimates presented in tables A1 through A4 suggest that the IV approach is reasonable. The end-of-life measures are strongly associated with inpatient spending (also see the partial F-statistics reported in tables 1 and 2) and only weakly, if at all, related to patients' illness severity. Moreover, we controlled for these characteristics in the regression, including measures of hospital spending on other types of patients. Reduced form estimates of associations between end-of-life measures and patients' thirty-day mortality are negative, statistically significant, and of plausible magnitude, except in the case of AMI patients. Nevertheless, we recognize the difficulty of implementing a valid instrumental variables procedure, and we cannot rule out the possibility that there may be some omitted variables affecting patients' mortality and failure-to-rescue that are correlated with the end-of-life measure. If so, then the IV estimates reflect the effect of inpatient spending and those characteristics affecting mortality and failure-to-rescue that are correlated with the end-of-life measures (e.g., quality).